Refine Search

Search Results -

Terms	Documents
L3 and geocod$	6

Database:

US Pre-Grant Publication Full-Text Database
US Patents Full-Text Database
US OCR Full-Text Database
EPO Abstracts Database
JPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Search:

Refine Search

Recall Text Clear Interrupt

Search History

DATE: Saturday, February 19, 2005 Printable Copy Create Case

Set Name side by side	**Query**	**Hit Count**	**Set Name** result set
	DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=OR		
L7	L3 and geocod$	6	L7
L6	L5 and geocod$	4	L6
L5	L3 and ("global positioning system" or "gps" or satellite)	423	L5
L4	L3 and ("global positioning system" or "gps" of satellite)	577	L4
L3	radials and centroids	2781	L3
L2	L1 and centroids	0	L2
L1	map near radials	63	L1

END OF SEARCH HISTORY

L7: Entry 1 of 2 File: USPT Aug 8, 2000

DOCUMENT-IDENTIFIER: US 6101496 A
TITLE: Ordered information geocoding method and apparatus

Brief Summary Text (2):
The background and the invention are best understood by defining certain terms including: geocoding, centroids, and street vectors/segments.

Brief Summary Text (4):
A centroid is a geographic center of an entire area, region, boundary, etc. for which the specific geographic area covers.

Brief Summary Text (5):
Street vectors are address ranges that are assigned to segments of individual streets. Street vectors are used in displays of digitized computer based street maps. Street vectors usually appear as left and right side address ranges. They are also used for geocoding a particular address to a particular street segment based on its point along the line segment. For example, the table below shows the address range on both sides of the street for one particular street segment of Main St.:

Brief Summary Text (8):
The georeferenced library is compiled from a number of varied sources including US Census address information and US Postal address information, along with Zip Code boundaries and other various sources of data containing geographic information and/or location geometry. If a raw data address cannot be matched exactly to a specific library street address (known as a "street level hit"), then an attempt is made to match the raw data address to an ever decreasing precision geographic hierarchy of point, line or region geography until a predetermined tolerance for an acceptable match is met. The geographic hierarchy to which a raw data record is finally assigned is also known as the "geocoding precision." Geocoding precision tells how closely the location assigned by the geocoding software matches the true location of the raw data. Current geocoding technology generally provides for two main types of precision: Street Level and Postal ZIP Centroid. Street Level precision is the placement of geocoded records at the street address. (See FIG. 1, record no. 1.) Street level precision attempts to geocode all records to the actual street address. In all likelihood, some matches may end up at a less precise location such as a ZIP centroid (ZIP+4, ZIP+2, or ZIP Code) or shape path (the shape of a street as defined by points that make up each segment of the street). A record is assigned or geocoded to the centroid of the shape path (S4--not listed in FIG. 1 as this is a rare occurrence) if the matching street address does not contain address ranges.

Brief Summary Text (9):
ZIP centroid precision places geocoded records at a postal record ZIP Code centroid. ZIP centroid precision matches a raw data record to the most precise ZIP Code it finds. The most precise postal match is one made to a ZIP+4 centroid. See FIG. 1, record no. 2. ZIP+4 is nearly as precise as a street level hit (street address). If a ZIP+4 centroid cannot be matched or does not exist, a match may then fall back to a ZIP+2 centroid (record no. 3) if available. The least accurate postal match is one made to a 5 digit ZIP centroid (record nos. 4, 5, 6.) If no street level or postal match can be found in the georeferenced library, then a

record remains ungeocoded (record nos. 7, 8, 9, 10). This can be the result of a lack of information in the georeferenced library (new building/development, address overlooked/not included, etc.) or a lack of information (missing address information, etc.) in the raw data records which are being geocoded.

Brief Summary Text (10):
One of the disadvantages of ZIP Code matching alone (without street address) is that current geocoding technology only examines the ZIP Code field when matching. If the ZIP Codes in the raw data records do not already have ZIP+4 values, then current geocoding technology will only match to the much larger area 5-digit ZIP Code centroids. Conversely, if you use Street Level precision, current geocoding technology will attempt to return street-level coordinates and will optionally fallback to the slightly less precise ZIP +4 coordinates. If the georeferenced library does not contain a full 9 digit ZIP Code (ZIP +4) x,y location for the raw data address, current geocoding technology will fallback on the less precise 5 digit ZIP coordinates.

Brief Summary Text (11):
As described above, another disadvantage of ZIP code matching is that ZIP+4 centroids may not exist at all and the only option is a fallback to the much larger area 5-digit ZIP Code centroid. An examination of current (January, 1998) ZIP+4 centroid availability bares out the problem of relying solely on ZIP+4 centroid placement when a specific street level address can not be found for a raw data record. FIG. 8 shows the breakdown of the ZIP+4 file for New York State. Fully two thirds of the centroids found in the file are not actually ZIP+4 centroids at all, but merely the less precise 5 digit ZIP or ZIP+2 centroids.

Brief Summary Text (21):
In the United States, the U.S. Census Bureau assigns street vectors. They are assigned during the decennial census by enumerators or "street canvassers" who do the actual census taking. Those address ranges are then compiled, digitized and otherwise made into street segments that contain address ranges or street vectors as described above. A compilation of those computer mapped streets for the entire U.S. is then made available for purchase through the Topologically Integrated Geographic Encoding and Referencing (TIGER) digital database.

Brief Summary Text (24):
The invention recognizes that there are a number of non-traditional data sources with geographically ordered information (OI). These non-traditional OI data sources include and are not limited to: tax property parcel records as maintained by state, county and municipal offices; insurance, disaster abatement, and fire code/regulatory records; various government records and privately held databases. The tax property parcel records are kept by state, county and municipal assessors offices for the maintenance of tax assessment, levy and property management. They offer unique OI. In most cases they are current, include new building developments, and offer a more comprehensive address database than traditional census and postal records. As such, OI records may not match addresses in traditional georeferenced libraries used in current geocoding technology. Therefore, it is not possible to assign precise x,y locations to those records that are not included in the traditional georeferenced library. That can pose a problem when geocoding a customer list in a new developments or in areas overlooked or not completely canvassed by the decennial census, for example. A georeferenced library based upon traditional (census and postal records) may not include precise street address coordinates for the new developments, etc. In such cases, the geocode precision will fall back to the less precise 5 digit ZIP code centroid found in the postal data portion of the georeferenced library. See FIG. 1 for samples of different types of geocoding precision. However, I have discovered a way of adding the OI information to the traditional database and for interpolating OI data to further enhance the precision of the georeference database.

Brief Summary Text (25):
The ordered information geography (OIG) algorithm process generates a much more precise x,y (z) coordinate placement at the Census Block centroid, Block Group centroid or other smaller area geography. By using the OI record identification keys (OIID), such as the property parcel identification number as assigned by the assessor, and then algorithmically processing them and including them in the georeferenced library, records are further geocoded with the OIID inherent geography. After geocoding in the traditional manner using existing geocoding technology, we assign locational coordinates to many of the OI records in a given area. We next use a series of select dialogues and programmatic queries to prepare those OI identification keys that are attached to the already geocoded records for greater location precsion assignment of less precise and ungeocoded records. We then assign a similar coordinate to the less precise and ungeocoded records based on similar or ranged and sorted on predetermined criteria OI identification keys. These additionally geocoded records are assigned to more precise centroids such as a census block centroid which can be the next best thing to actual rooftop or street level geocoding.

Drawing Description Text (2):
FIG. 1a is a table showing examples of prior art geocoded records with data fields including centroids of different precision including: street level hits, ZIP+4 hits, ZIP+2 hits, ZIP hits, and ungeocoded records;

Drawing Description Text (4):
FIG. 2 is schematic diagram of computer programmed to carry out the geocoding of the invention;

Drawing Description Text (5):
FIG. 3 is a high level flow diagram of a computer program for carrying out the invention;

Drawing Description Text (6):
FIG. 4 is a more detailed Wanier--Orr diagram of the portion of the computer program that sorts the records by the precision of the centroids and assigns greater location precision to less precise geocoded OI records and ungeocoded OI records which are then inserted to enhance the georeferenced library. FIG. 4 also describes the method to enhance the street segment address library;

Drawing Description Text (7):
FIGS. 5, 6, and 7 are diagrams of examples of computer records before and after operation of the program;

Drawing Description Text (8):
FIG. 8 is a table of the ZIP+4 centroids in New York State;

Drawing Description Text (9):
FIG. 9 is a graph of the changes in centroid location as a result of running the program;

Drawing Description Text (10):
FIG. 10 is a table of centroids taken before and after running the computer program;

Detailed Description Text (7):
The most precise (usually street level) geocoded records are assigned to the highest possible precision small area geometry for the particular geocoded area. In the United States, such records are usually the Census Bureaus TIGER records. They provide digital coverage of approximately seven million Federal Information Process Standard (FIPS) blocks whose individual borders represent the street segments found in computer cartographic street display and address products. In urban areas, FIPS

blocks are often the smallest digital area geometry available. It usually corresponds to an actual city block. The goal of this invention is to geocode to the center of the available highest precision small area geography, which in the US is usually the FIPS Block centroid.. Another precise small area geometry is the ZIP+4 coverage for the United States. However, ZIP+4 coverage is spotty at best and geocoders often fall back to the much less precise 5 digit ZIP. Consider, for example, the ZIP codes used in New York State. As shown in FIG. 8, there are over three million ZIP code centroids in New York. However, less than half are ZIP+4 and more than half are simple 5-digit ZIP codes. Any geocoded New York information based on ZIP codes will have very limited precision because more than half of the ZIP+4 x,y locations may fall back to the ZIP centroid. However, the OIG process of the invention improves precision by assigning more precise locations to many of the existing (and future) 5, 7 and 9-digit ZIP codes.

Detailed Description Text (8):
The invention increases the total number of raw data records that are geocoded by using a new methodology in combination with current geocoding technology. With reference to FIGS. 2 and 3, the invention comprises a computer with the inventive program, the program stored on a disc, and a series of steps for operating a computer to improve an existing geocoded library. Its novel features include geocoding OI records using current technology for various location precision assignments; merging high precision results with varying geographies (attaching precise geography such as block regions, etc.) as part of the geocode process; interpolating geocoded OI individual record identification keys and their sequential, alphanumeric or other location component in order to assign more precise locations to records and to an enhanced highly precise or higher precision locations; merging these enhanced precision non-traditional OI location records with the georeferenced library in order to create a larger georeferenced library for improved geocoding.

Detailed Description Text (9):
FIG. 2 shows a computer system 10 including a central processing unit (CPU) 14. The CPU 14 is well known. It may comprise a personal computer CPU or a large, main frame CPU. It has one or more execution units including one or more arithmetic logic units and one or more floating point units. A user has an input device, such as a keyboard 12 to enter data into the computer 14. Other input devices may be used, including and not limited to a mouse. A memory 20 holds programs and one or more databases. Database 21 is an existing GL of first records. Database 22 has OI data in a second set of records. The CPU 14 compares the records in database 21 to those in the database 22 to generate a third set of records 23.

Detailed Description Text (12):
OIG 1. After initial geocoding and processing of OI records, the computer selects all from the S5, S4 and S3 (high precision [HP] street level geocoded records). In this case, the assigned precision geography (APG) is the 15 digit census FIPS block code. In this case, the APG is the 15 digit census FIPS block code. Each APG usually contains at least 2 or more geocoded point records.

Detailed Description Text (18):
FIG. 5 shows an example of how the invention improves a typical ZIP, centroid hit to a street level hit and thereby improves the geocoding by relocating the position of the address more than two miles closer to its actual location. Campbell Block is geocoded for addresses 2-6 and 8-34. It is not geocoded for 58 Campbell, which is assigned to its nearest ZIP centroid, the star shown in the lower middle of 5. However, after the existing library is enhanced with the OI taken from the tax roles, 58 Campbell is relocated to the upper left hand comer of FIG. 5. The change in precision is 2.16148 miles closer to the actual location.

Detailed Description Text (20):
Traditional geocoding requires a user have intimate knowledge of the Albany area

and the particular address she/he is attempting to geocode to know how to separate the Albany Hillcrest addresses from the Westmere Hillcrest addresses. In most cases, the user is not familiar with the geography of the area he/she is geocoding and the only option is to place a geocoded record to the 5 digit zip centroid precision in cases where the same street name is repeated within the same zip code. A centroid location 50 is given for 14 Hillcrest addresses. The OIG process places the location of the raw data record in the block centroid or interpolates a street level segment closest to the correct similarly named street. This information along with address specifics, etc. is stored in the enhanced georeferenced library lending more "intelligence" for future geocoding runs. As a result, the fourteen addresses are relocated to more precise locations 51-54.

Detailed Description Text (22):
A direct benefit from running the OIG process to increase the number of pinpointed x,y (z) addresses in the geocoding georeferenced library is the ability to interpolate from these addresses the near or exact location of new street segments containing the vector of these address ranges. The street segment product, as described, is often used to display information through various computer cartographic or presentation graphics. When an individual wishes to visualize a geocoded record set, these records are placed on their corresponding street vector and displayed upon various vector and/or raster coverages. Using existing street segment coverages, we can extend segments using the high precision OIG location points as determinants in assigning which vector to add to as well as direction and size of the new street segment.

Detailed Description Text (23):
FIG. 7 shows how the invention locates new segments of streets in the existing database. Before the invention is used a Berne Altmont street segment has geocoded address 1900-2350 with a number of street level hits and with addresses 2371 and 2365 Berne Altmont located at ZIP centroid 60. Since 2371 and 2365 are not included in the street level hits, the database defaults to the ZIP centroid. The OI data indicates that there is a new segment of Berne Altmont with addresses 2351-2399. After processing the existing database with the relevant OI data, the new segment of 2351-2399 is added to the existing segment and the 2371 and 2365 Berne Altmnont addresses are relocated to the new segment based on the proximity of the OIG assigned point to the ending existing street segment, at locations 62, 63, respectively. In addition, the OIID of both the high precision geocoded OI records and the positive matched records of the less precise OI records which have been assigned greater precision through the OIG process can be used to create a topological structure, giving direction and adjacency for creation of new street segments/vectors in the SSAD. This topological structure can be interpolated from the inherent geographic information contained in the OIID once actual x,y location is assigned to sequentially proximate records using the OIG process, allowing for more precise placement of additional address ranges represented as street segments and/or points on a map. In addition to the invention creating new street segments and associated vectors from comprehensive address sources such as tax property parcel records, the high precision OI record location points and low precision records assigned a greater location precision through the OIG process can be used as "point vectors" or address ranges condensed to a single x,y coordinate point. This is in essence a way of adding entirely new "streets" to a street display or addressed products. Although represented graphically by a point rather than a street line, these provide higher geocoding hit rates when geocoding is performed through various proprietary software rather than against a georeferenced library in a geocoding engine.

Detailed Description Text (26):
A sample of an OIG process was run on Albany County in New York State. FIGS. 9 and 10 demonstrate how the invention makes more precise x,y location assignments to S1 or 5 digit Zip centroid location assignments than location assignments available through traditional geocoding methods. For various reasons (missing address

information, same name streets in same small area geography and indiscernible match conditions) traditional geocoding returned a S1 (ZIP) precision level hit. Of 15 cases, 13 original S1 hits were more precisely located under the OIG process. This analysis was possible because a corresponding street segment was available at or near the true Cartesian location. See FIG. 11. A complete breakdown of the individual address assignments for each S1 assignment or group of assignments is shown in FIG. 10. FIG. 12 shows that overall location accuracy performance of the OIG process in comparison to S1 assignment under traditional geocoding methods is 2.3 times greater for this example.

CLAIMS:

1. A method for improving a geocoded database comprising the steps of:

comparing a first set of geocoded database records to second set of records containing inherent geographic information,

said first set of records each comprising a first number of data fields including data representing an identification of a geographic location

corresponding to the record and data representing one of two or more geographic centroids representative of geographic areas including the location, said centroids from a centroid with highest precision to a centroid with lowest precision;

said second set of records comprising inherent geographically ordered data fields where said data represents a unique identification of a geographic location and the proximity of one record of one location to other records at other geographic locations and one or more data fields corresponding to the data fields of the records in the first set;

generating a plurality of matches where a record in the first set has a data field that matches a data field of a record in the second set;

sorting the matched sets by the centroids of the first set of records;

selecting matched sets with the highest precision centroids;

adding the geographically ordered data fields of the second set to the records matched in the first set to generate a third set of records.

2. The method of claim 1 comprising the further step of comparing the third set of records to the second set of records to identify records in the second set that are geographically proximate to one or more records in the third set;

changing the centroid of the identified records of the second set to correspond to the centroid of the most proximate record in the third set; and

adding the geographically ordered data fields of the second set to the most proximate records of the third set.

3. The method of claim 1 wherein the centroids comprise at least four centroids of different precision.

4. The method of claim 3 wherein the centroids comprise street level, ZIP+4, ZIP+2, and ZIP.

5. The method of claim 1 wherein the step of selecting comprises selecting the matched sets for the highest and the second highest precision centroids.

6. The method of claim 5 wherein the highest and second highest precision centroids

are street level and ZIP+4.

7. The method of claim 6 wherein the step of selecting matched sets comprises selecting matches with the highest and second highest precision centroids to create a third set of records;

and further comprising comparing the third set of records to the second set of records to identify records in the second set that are geographically proximate to one or more records in the third set;

changing the centroid of the identified records of the second set to correspond to the centroid of the most proximate record in the third set; and

adding the geographically ordered data fields of the second set to the most proximate records of the third set.

12. A computer program stored on a disc and comprising a program for geocoding a database comprising the steps of:

comparing a first set of geocoded database records to second set of geographically ordered records,

said first set of records each comprising a first number of data fields including data representing an identification of a geographic location corresponding to the record and data representing one of two or more geographic centroids representative of geographic areas including the location, said centroids from a centroid with highest precision to a centroid with lowest precision;

said second set of records comprising inherent geographically ordered data fields where said data represents a unique identification of a geographic location and the proximity of one record of one location to other records at other geographic locations and one or more data fields corresponding to the data fields of the records in the first set;

generating a plurality of matches where a record in the first set has a data field that matches a data field of a record in the second set;

sorting the matched sets by the centroids of the first set of records;

selecting matched sets with the highest precision centroids;

adding the geographically ordered data fields of the second set to the records matched in the first set to generate a third set of records.

13. The computer program of claim 12 comprising the further step of comparing the third set of records to the second set of records to identify records in the second set that are geographically proximate to one or more records in the third set;

changing the centroid of the identified records of the second set to correspond to the centroid of the most proximate record in the third set; and

adding the geographically ordered data fields of the second set to the most proximate records of the third set.

14. The computer program of claim 12 wherein the centroids comprise at least four centroids of different precision.

15. The computer program of claim 14 wherein the centroids comprise street level, ZIP+4, ZIP+2, and ZIP.

16. The computer program of claim 12 wherein the step of selecting comprises selecting the matched sets for the highest and the second highest precision centroids.

17. The computer program of claim 16 wherein the highest and second highest precision centroids are street level and ZIP+4.

18. The computer program of claim 17 wherein the step of selecting matched sets comprises selecting matches with the highest and second highest precision centroids to create a third set of records;

and further comprising comparing the third set of records to the second set of records to identify records in the second set that are geographically proximate to one or more records in the third set;

changing the centroid of the identified records of the second set to correspond to the centroid of the most proximate record in the third set; and

adding the geographically ordered data fields of the second set to the most proximate records of the third set.

19. The computer program of claim 12 wherein the records in both sets comprise data fields for the street address of the records and the matching step comprises comparing the street address data fields of records in the first set to the street address data fields of records in the second set.

20. The computer program of claim 12 comprising the further step of mapping further data to the third set of records.

21. A computer for geocoding a database comprising:

a memory for holding a first set of geocoded database records and a second set of geographically ordered records;

means comparing the first set of geocoded database records to the second set of geographically ordered records,

said first set of records each comprising a first number of data fields including data representing an identification of a geographic location corresponding to the record and data representing one of two or more geographic centroids representative of geographic areas including the location, said centroids from a centroid with highest precision to a centroid with lowest precision;

said second set of records comprising inherent geographically ordered data fields where said data represents a unique identification of a geographic location and the proximity of one record of one location to other records at other geographic locations and one or more data fields corresponding to the data fields of the records in the first set;

means for generating a plurality of matches where a record in the first set has a data field that matches a data field of a record in the second set;

means for sorting the matched sets by the centroids of the first set of records;

means for selecting matched sets with the highest precision centroids;

means for adding the geographically ordered data fields of the second set to the records matched in the first set to generate a third set of records.

22. The computer of claim 21 further comprising means for comparing the third set

of records to the second set of records to identify records in the second set that are geographically proximate to one or more records in the third set;

means for changing the centroid of the identified records of the second set to correspond to the centroid of the most proximate record in the third set; and

means for adding the geographically ordered data fields of the second set to the most proximate records of the third set.

23. The computer of claim 21 wherein the centroids comprise at least four centroids of different precision.

24. The computer of claim 23 wherein the centroids comprise street level, ZIP+4, ZIP+2, and ZIP.

25. The computer of claim 21 wherein the means for selecting comprises means for selecting the matched sets for the highest and the second highest precision centroids.

26. The computer of claim 21 wherein the highest and second highest precision centroids are street level and ZIP+4.

27. The computer of claim 26 wherein the means for selecting matched sets comprises means for selecting matches with the highest and second highest precision centroids to create a third set of records;

means for comparing the third set of records to the second set of records to identify records in the second set that are geographically proximate to one or more records in the third set;

means for changing the centroid of the identified records of the second set to correspond to the centroid of the most proximate record in the third set; and

adding the geographically ordered data fields of the second set to the most proximate records of the third set.

28. The computer of claim 21 wherein the records in both sets comprise data fields for the street address of the records and the means for matching comprises means for comparing the street address data fields of records in the first set to the street address data fields of records in the second set.

29. The computer of claim 21 further comprising means for mapping further data to the third set of records.

Previous Doc Next Doc Go to Doc#